**WHAT IS CLAIMED IS:**

1. A method for generating a natural language understanding model, comprising:

collecting a plurality of utterances;

generating a plurality of call types each having utterances selected from said collection of utterances, said utterances used to generate said plurality of call types representing a first set of utterances which is a subset of said collection of utterances;

generating a first natural language understanding model using call type information contained within said first set of utterances;

testing said first natural language understanding model;

modifying said plurality of call types based on said testing; and

generating a second natural language understanding model using said modified plurality of call types.

2. The method of claim 1, further comprising generating an annotation guide using a second set of utterances which is a subset of said first set of utterances.

3. The method of claim 1, further comprising generating call type data using at least one of data clustering, relevance feedback, string searching, data mining, and active learning tools.

4. The method of claim 3, wherein said call type data is generated using a graphical user interface.

5. The method of claim 1, wherein said first natural language understanding model is trained using a first text file containing utterances contained within said first set of utterances and a second text file containing call types assigned to said utterances in said first text file.

6. The method of claim 1, wherein said natural language understanding model is tested using a subset of said first set of utterances.

7. The method of claim 1, wherein said plurality of call types are modified using a graphical user interface.

8. The method of claim 1, wherein said first natural language understanding model is created prior to an annotation guide.

9. A spoken dialog system, comprising:

a natural language understanding model trained using a method comprising:

collecting a plurality of utterances;

generating a plurality of call types each having utterances selected from said collection of utterances, said utterances used to generate said plurality of call types representing a first set of utterances which is a subset of said collection of utterances;

generating a first natural language understanding model using call type information contained within said first set of utterances;

testing said first natural language understanding model;

modifying said plurality of call types based on said testing; and

generating a second natural language understanding model using said modified plurality of call types.

10. The system of claim 9, wherein the method further comprises generating an annotation guide using a second set of utterances which is a subset of said first set of utterances.

11. The system of claim 9, wherein said call type data is generated using at least on one of data clustering, relevance feedback, string searching, data mining, and active learning tools.

12. The system of claim 11, wherein said call type data is generated using a graphical user interface.

13. The system of claim 9, wherein said natural language understanding model is trained using a first text file containing utterances contained within said first set of utterances and a second text file containing call types assigned to said utterances in said first text file.

14. The system of claim 9, wherein said natural language understanding model is tested using a subset of said first set of utterances.

15. The system of claim 9, wherein said plurality of call types are modified using a graphical user interface.

16. The system of claim 9, wherein said first natural language understanding model is created prior to an annotation guide.

17. A method for generating a natural language understanding model, comprising:

collecting a plurality of utterances;

generating a plurality of call types each having utterances selected from said collection of utterances, said utterances used to generate said plurality of call types representing a subset of said collection of utterances; and

generating a natural language understanding model using call type information contained within said subset of utterances, wherein said natural language understanding model is generated prior to receipt of manually labeled utterance data.

18. The method of claim 17, wherein said manually labeled utterance data is generated using an annotation guide that is created using a portion of said subset of utterances.

19. The method of claim 17, wherein said natural language understanding model is generated using a first text file containing utterances contained within said subset of utterances and a second text file containing call types assigned to said utterances in said first text file.

20. The method of claim 17, wherein said natural language understanding model is tested using a second subset of said collection of utterances.